From Question-Answering to Information-Seeking Dialogs

Jerry R. Hobbs
SRI International

Questions are usually part of a larger information-seeking activity, and we will not have adequate question-answering systems until the information-seeking activity itself is modelled. Analysts typically have a large issue they want to elucidate, and the questions asked are in the service of that need; from the answers taken together the analyst develops a position on the issue of interest. The position bears only an inferential relation with the answers to the questions. For the "question-answering" system to participate fully in the information-seeking activity, it must be able to aid and track this inference process.

SRI International and USC Information Sciences Institute are collaboratively pursuing research aimed at developing the technology and the corresponding "information-seeking" system that will enable computers to be full partners in an information-seeking dialog. The key ideas are to

- Do a logical analysis or decomposition of questions into component sub-questions, using our inference engine SNARK, with several currently available knowledge bases, together with knowledge bases to be built or augmented in connection with this project
- Use the inference engine with high-level rules to determine the likely place of a query in a larger information-seeking activity, and use that for contextually dependent interpretation of questions
- Use the component questions to drive the subsequent dialog between the analyst and the system, to track the state of the dialog, to elicit elaborations, revisions, and clarifications, and to suggest modified questions whose answers would be more informative

- Determine the answers by bottoming out, via SNARK's procedural attachment capability, in a variety of Web resources, including Teknowledge's search engine for DAML-encoded Web sites, an information extraction system called TextPro developed at SRI, and questions back to the analyst and communications with other experts

- Use the inferential structure of the inquiry thus discovered to recompose the answers found for the subquestions into an answer for the question as a whole, and formulate and present this answer in a manner structured by the recomposition.

Our approach is to translate queries into a logical form using the GEMINI parser and grammar, and then attempt to prove that logical form, or as much of it as possible, using the SNARK inference engine on a knowledge base of axioms of several sorts. The motivating idea behind this approach is that the question is a high-level description of an information need, and by proving it we supply the information need that is described. In the course of the proof we resolve indeterminacies inherent in the question, and we translate between the vocabulary of the question and the ontologies of the available information sources, which are then invoked to supply parts of the desired answer. The structured analysis of the query that led to the answer is then used to reconstruct the answer for presentation to the user. It can also be used to structure the dialog with the user.

Realizing this vision will entail research in ten different areas.

- Parsing Queries: The GEMINI parser and grammar are already highly developed for question-answering applications. We have had to augment the grammar only in minor ways and do a small amount of development to produce the kinds of logical forms the inference engine will need.

- Resolving Indeterminacies: The logical forms produced by GEMINI contain certain indeterminacies. General predicates need to be given specific interpretations, and metonymies and coference must be resolved, for example. In previous research we have shown that solutions to these problems often arise as a side effect of proving the logical form of the sentence. In this project, we will make use of and extend that work, experimenting with the axiomatizations required to resolve typical indeterminacies in questions.

- Decomposing Questions: The answers to questions, especially complex questions, are often not to be found in a single place, but must be composed from information found in several places. The questions must be decomposed into their parts, as determined by where the information will be found. Axioms in the knowledge base direct the inference process toward different available information sources for the answers to different parts of questions. Constructing these axioms in a systematic fashion has been a primary focus of our research on the AQUAINT project.

- Articulating with Resource Ontologies: There must be rules that mediate between the ontologies inplicit in the questions and the ontologies implicit or explicit in the information sources. These rules may be thought of as "articulation" axioms, in the sense that they articulate, or fit together, two different perspectives on a subject area. In our AQUAINT project, we make various resources available to our AQUAINT system, and in doing this, one of the principal tasks is encoding the articulation axioms that make communication with the resources possible.

- Invoking Resources: In our project, we connect the AQUAINT system with a number of diverse resources, and for each of these, procedural attachment axioms are written.

- Encoding Axioms: One of the principal tasks of this project has been to encode the axioms required for decomposing information needs, articulating with resource ontologies, and invoking resources.

- Filtering Responses from Resources: Very often the analyst will require very specific information from a source that can only be queried on general criteria. In these cases the inference engine filters the answers on the basis of the more specific constraints.

- Structuring Answers: In our approach, a question is decomposed into subquestions, and these subquestions are answered by various resources. It is then necessary to recompose the information found into a coherent answer to the analyst's question. Because the inference process proceeds hierarchically, the hierarchical nature of this process can be used to to construct a hierarchically organized answer. This structure will also make drill-down and explanation easier. We will conduct research to determine how best to use the structure of the inferences leading to answers to structure those answers.

- Structuring Dialog: The structure of the inference process can also be used in carrying on dialogs with an analyst seeking information. It provides the context necessary for interpreting subsequent queries in a dialog and for making requests for clarification and elaboration. We will investigate this use of the inference process for tracking and managing the state of the dialog.

- Measuring Reliability: To a limited extent in Phase I of the project, we will investigate the use of a quantitative scheme developed previously for measuring the reliability of answers, based on the reliability of defeasible axioms and information sources that are not necessarily perfect. We expect this to be a more major focus of research in subsequent phases of the project.

We believe these ideas represent a large part of the long-term solution to the information-seeking problem. But in addition we are finding that because the resources we use are already highly developed, significant capabilities are already being achieved.

The two principal achievements of the last six months have been the incorporation of an information extraction engine and a DAML search engine into our system.

The information extraction engine is TextPro, developed at SRI and used most recently for the ACE program evaluations. It has been consistently one of the leaders in information extraction evaluations. The problem we have faced is that the output produced by TextPro is annotated text and the SNARK theorem prover must invoke TextPro and use its results in the form of logical expressions. Thus, significant work had to be done to translate between the two. The advantage of using TextPro as a component of the larger system, as opposed to having it be a stand-alone system, is that it can be used to answer subquestions of the questions asked, and the answers can be integrated with information obtained from other resources.

In October we learned about the ASCS search engine developed by Teknowledge to search the entire collection of web pages expressed in DAML, DARPA's candidate for the basis of the "Semantic Web". We immediately began discussing with them the possibility of using ASCS as a resource in our

AQUAINT system. Their interface required logical expressions very much like the logical expressions used in SNARK. It thus has proved very straightforward to incorporate it. This gives our AQUAINT system access to the soon-to-be exponentially growing content on the Semantic Web, and it gives ASCS a natural language interface, thus making it accessible to vastly many more people than before.

In addition we have been enhancing our ontologies of space and time, and have implemented a rudimentary ontology of agents and actions.